********* **DRAFT FOR COORDINATION** *********

JEDMICS Compact Disk Engineering Data Exchange (CDEX)

Addendum

17 January 1997

Introduction

On 1 October 1996, the JEDMICS Program Office issued a document titled JEDMICS Compact Disk Data Exchange (CDEX) File Structure and Format and has been implementing the JEDMICS CDEX interface at a number of sites. This document described the data structure and content model for input and output of data with JEDMICS via compact disk, provided basic data requirements, and provided an example of an on disk structure. The CDEX logical structure provides a highly flexible approach that is easily adapted to meet an organization's business process. This flexibility is possible because of a set of relationships and principles that insure integrity of the data structure while allowing sites to adapt and add additional information.

Some confusion has been generated among the user community because the **sample** disk structure provided one implementation of the structure and contains both required and optional directories or files. As stated in paragraph 1.1, Figure 1 is sample structure.

This purpose of this addendum is to provide an information that further clarifies the referenced document by discussing the essential relationships and principles inherent in the CDEX approach thus maintaining compatibility between DOD wide JEDMICS sites while providing flexibility to adapt to the numerous business processes that use technical repository data.

Reference Paragraph 1.1, JEDMICS Process.

The sample structure (Figure 1) is a sample that contains both required and optional directories. Sites may adapt the CDEX structure to their business process by adding additional directories/files to the structure in order to accommodate site specific requirements; provided that the essential logical structure described herein is retained. The essential required elements and relationships are described below.

Reference Paragraph 2.0. Metadata/Index Data File

The file name INDEX.DLF and metadata file term or used interchangeable throughout this document to refer to index metadata or the computer file that defines it.

The CDEX approach consists of a logical data structure where a volume identifier reference (i.e. VOLUMEID.TXT, JOBS_Contained Field) points to a metadata file, which in turn points to all data files that make up the dataset. This approach provides a logical relationship of the data with minimum constraints imposed by the physical media type or the physical file structure of the media. The fundamental principles are as follows:

A. The Metadata/Index Data File (i.e. INDEX.DLF) is a set of index data that describes the data files and provides a pointer to the location of the data files.

B. Each CDEX compact disk must have<u>at least one</u> file named INDEX.DLF that contains the index metadata associated with the data files. The default location of the INDEX.DLF file is the root directory.

C. However, a CDEX disk<u>may have several</u> INDEX.DLF files.
In the case of multiple packages (e.g. parts, NSNs or multiple solicitations) stacked on a disk, an entry is made in the VOLUMEID.TXT file (at the root directory) to point to the directory location of the appropriate INDEX.DLF files. This object oriented approach allows for identifying all packages on the disk, linking the top level identification of the package to the metadata for the package, and linking each metadata record to the corresponding data file.

D. An INDEX.DLF file applies to all data files in the current directory or its sub-directories. Conversely, all data files that are contained in the current directory (or a sub-directory of the current directory) will have an appropriate entry in the INDEX.DLF file.

E. The CDEX structure was designed to be flexible enough to adapt to many business processes that must exchange repository technical data. To accomplish this goal, the organization producing the data must be able to add additional data files to the structure. This can be accomplished by adding any data file to the dataset and adding the appropriate entry to the metadata file. In other words, add the data file to the set of data anywhere in the structure and insert an entry into the INDEX.DLF file. Multiple datasets (e.g. TDPs) can be contained by replicating the data structure and adding an entry to the VOLUMEID.TXT file at the root level.

JEDMICS Data File Index Structure (DFIS) Format and Structure Table:
The DFIS Format and Structure provides a common metadata structure for input and output interface with the JEDMICS. This structure contains all the necessary index data to populate the JEDMICS database upon delivery of data to the system interface and provides all available JEDMICS output data elements. Therefore, it is imperative that required data fields be populated IAW

the referenced table to insure that input data can be properly inserted into JEDMICS. The fundamental principles of this approach are as follows:

   A. Each INDEX.DLF file contains one record for each file in the set of data files (i.e. dataset). Each record is separated by a carriage return and line feed character (Hex 0A0D) per the DFIS Table. Each field of data is separated by a pipe character (|) thus allowing the data to be created, read or modified by humans and parsed by computers.

   B. The INDEX.DLF file must consist of alpha-numeric characters IAW ANSI Standards. This principle allows the data to be easily prepared by humans (or computer interface) and easily parsed by computer interface programs. This approach provides a capability for input data to be prepared in a number of ways ranging from using a simple text editor to a computer interface with the data developer's system. This approach allows even a many small business with a PC to prepare data for exchange with the government and a allows major primes to easily interface the structure with their data system.

   C. The index metadata must be in the format per DFIS Table. This common data format and structure insures that all JEDMICS systems can output and input data from another JEDMICS system and insures that external systems interoperate with JEDMICS.

   D. Required and option fields are defined in the DFIS Table notes. The required fields are based on the requirement to populate database key fields upon input to JEDMICS. The JEDMICS output process provides a complete set of available data. By providing a complete set of data available from JEDMICS, the numerous uses/applications of this data can select that data which is pertinent to the business process. The DFIS format does not dictate the values of the data fields. The format shows the allowable field sizes based on the JEDMICS database. Allowable field values are shown in the JEDMICS 2.5 User Guide.

   E. The DFIS Table Data, File Name and Path Fields entries <u>must</u> be entered thus providing a logical pointer to every file in the dataset. The logical pointer approach means that the data files may exist anywhere on the disk provided the INDEX.DLF file contains the pointer (i.e. file name and path name field) to file. This approach divorces the logical data structure from the physical media structure limitations and allows the data to be transferred between any two hierarchical file systems.

In other words, because all computer systems support some type of hierarchical file system, the data structure can continue to exist when the data is moved from optical storage to a hard drive, to a compact disk, to telecommunications, to cartridge disk, and media not yet widely available.

F. The only limitations on data file naming convention (i.e. the name of the file for the image data) is the ISO-9660 constraint of a unique name with a maximum of 8 characters. Because the INDEX.DLF metadata file contains the filename and path to the data file, any value can be given to the actual data file name. The naming convention used in the MIL-STD-1840 specification is acceptable but not mandatory.

G. The Data Format Type field must be populated. This field indicates the type of data file (e.g. Raster, vector, text, document). As DOD is transitioning to contractor format data, it is imperative that the data format type be identified in the metadata. Historically, the governments technical data repository has managed only raster images. However, the JEDMICS allows management of multiple types of data. The allowable data format types are show in the JEDMICS 2.5 User Guide.

H. Since a compact disk may contain thousands of data files (e.g. images), the JEDMICS output process places the data files in a directory named IMAGES as a convenience. Because the INDEX.DLF file contains a logical pointer to the data file, the IMAGES directory may contain sub-directories to further organize the data. In essence, the data file may be located anywhere on the disk and the logical pointer used to located the data file. The logical pointer approach may be easily expanded to locate data files on multiple disk volumes.

Reference Paragraph 2.3 Disk Marking/Labeling:
Compact disks produced from JEDMICS will be labeled with identification information, security information and distribution rights information IAW DOD guidance e.g. MIL-STD-1806 (Digital Data Marking), MIL-HDBK-9660 and DFARS. The disk labeling application used with JEDMICS CDEX is extremely flexible thus allowing sites to customize with site specific logos and select a variety of data from the VOLUMEID.TXT file or default templates.

Reference Paragraph 3.0 Compact Disk Data Structure
As described in previous paragraphs the JEDMICS CDEX approach implements a logical data structure where the VOLUMEID.TXT file at the root level identifies the datasets on the disk by identifying the directory of the metadata files. The metadata file (i.e. INDEX.DLF) identifies the data files that make up the dataset by providing a pointer to each file.

The referenced JEDMICS CDEX document example disk structure shows several directories such as REQUEST, RFQ, TDPLS, IMAGES, NORS, MODELS. These directories are optional and are presented as one example of a physical file structure that can exist within the logical data structure. Data can be stored within any of these directories and the appropriate entry added to the INDEX.DLF file thereby creating a physical data structure that meets the users

business process requirements while retaining a logical data structure that enables exchange of data.

Reference Paragraph 3.1 Volume Identification File.

Per MIL-HDBK-9660 and other DOD guidance, each electronic media must contain a file with information identifying the organization point of contact who created the disk, security information, distribution information, and other reference information. The VOLUMEID.TXT file sample shows the minimum set of information referenced by the above guidance. The fundamental principles are as follows:

A. The purpose of the VOLUMEID.TXT file is to provide identification information for the electronic media as required by DOD guidance. The README.TXT file should be used to provide general information, instructions and general comments.

B. The VOLUMEID.TXT file must consist of alpha-numeric characters IAW ANSI Standards thus allowing the file to be easily read by humans. The JEDMICS output process generates the data in a tuple format that consists of a constant text string followed by a colon (:) character, followed by a text string. This approach allows the file to be easily created or read by both humans and computers.
The JEDMICS output interface creates this file by combining data specific values with a site defined template.

C. All the data fields shown in the example may not apply to all organizations (or a specific use of the structure) and may be blank if the field does not apply.

D. The Jobs_Contained field must be populated with the directory path name where a metadata file (INDEX.DLF) is located. When multiple datasets are stacked within the structure, an entry must be provided for each dataset. In other words, one Jobs_Contained entry must be present for each INDEX.DLF file present on the disk.

Reference Paragraph 4.0 Compact Disk Mastering:

The JEDMICS output process will produce compact disks compliant with ISO-9660 (International Standards Organization Specification 9660) and ISO-10149 Mode 1 (Yellow Book) physical format with data files provided IAW CALS Handbook 59 (MIL-HDBK-59). The rational for this approach is as follows:

A. Considering that computer technology is changing rapidly with new formats for compact disk being introduced at a rapid rate; and the fact that the intended user of the CDEX is a broad sector from small business to major

corporation; a common denominator compact disk physical format was selected to promote the widest possible acceptance. The logical data structure approach described above allows DOD to adapt newer technology as it becomes available.

      B. The ISO-9660 format is the only format supported by all computer manufacturers including Intel/Microsoft Based Personal Computers (PCs), Apple Computers, and UNIX based computers. While the government provided software included on the disk can only be used on a PC, the data files may be accessed and used by numerous commercial viewer software products. The ISO-10149 defines the physical media for the ISO-9660 data format.

      C. The JEDMICS output image data files are provided in CALS compliant format IAW MIL-HDBK-59 CALS Standards (such as MIL-D-28001, MIL-D-28002 Type 1 or Type 4, and others). The CALS compliant approach allows the end user to easily access and transfer the data files to his/her local computer system and provides the fundamental enabler for contracted data delivery to the government. Considering that; numerous contracts already exist where digital data files in CALS Compliant format must be delivered to the government; and future government contracts will require digital delivery of data files (per DOD 5000); it is imperative that data files be exchanged in CALS compliant format.